Supplementary Information:

The evolution of sensory divergence in the context of limited gene

flow in the bumblebee bat

Sébastien J. Puechmaille, Meriadeg Ar Gouilh, Piyathip Piyapan, Medhi Yokubol, Khin

Mie Mie, Paul J. Bates, Chutamas Satasook, Tin Nwe, Si Si Hla Bu, Iain J. Mackie, Eric

J. Petit and Emma C. Teeling

* Corresponding authors: s.puechmaille@gmail.com; emma.teeling@ucd.ie

This Supplementary Information includes:

- Supplementary Figures S1-S5,

- Supplementary Tables S1-S9,

- Supplementary Notes 1-2,

- Supplementary Methods,

- Supplementary References.
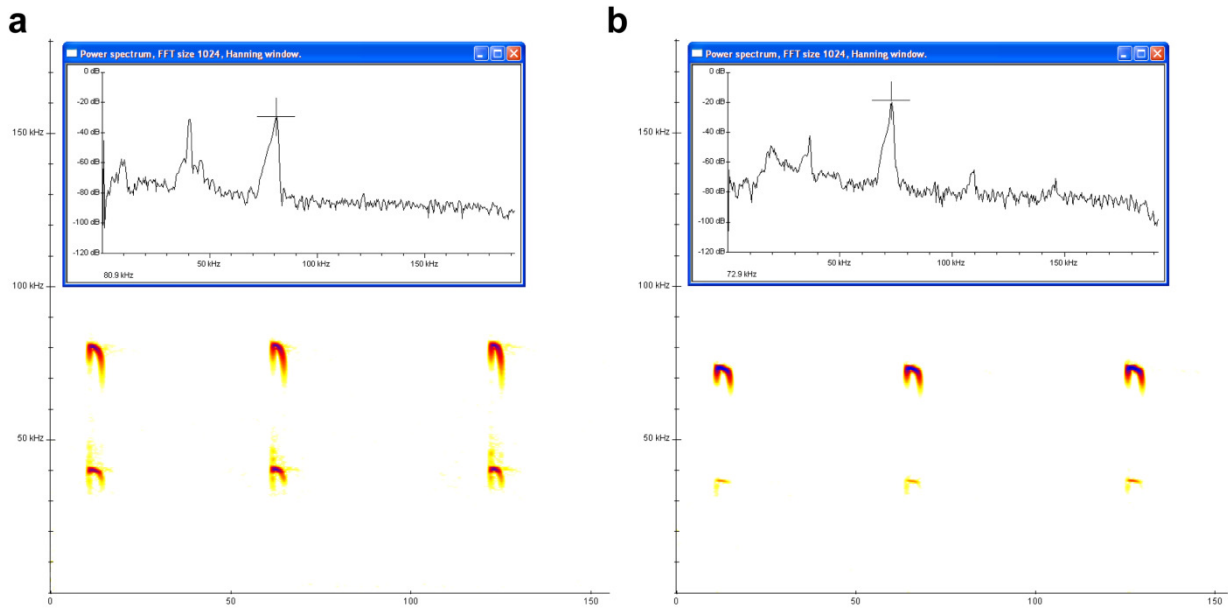
Supplementary Figure S1 | Echolocation calls of *Craseonycteris thonglongyai* in Thailand and Myanmar. Spectrogram (bottom) and power spectrum (top) showing the difference in frequency between *C. thonglongyai* echolocation calls **(a)** in Myanmar (left, peak frequency=80.9 kHz, colony M3) and **(b)** Thailand (right, peak frequency=72.9 kHz, colony A). Note that the X-axis (peak frequency in kHz) and Y-axis (time in ms) scales are identical.
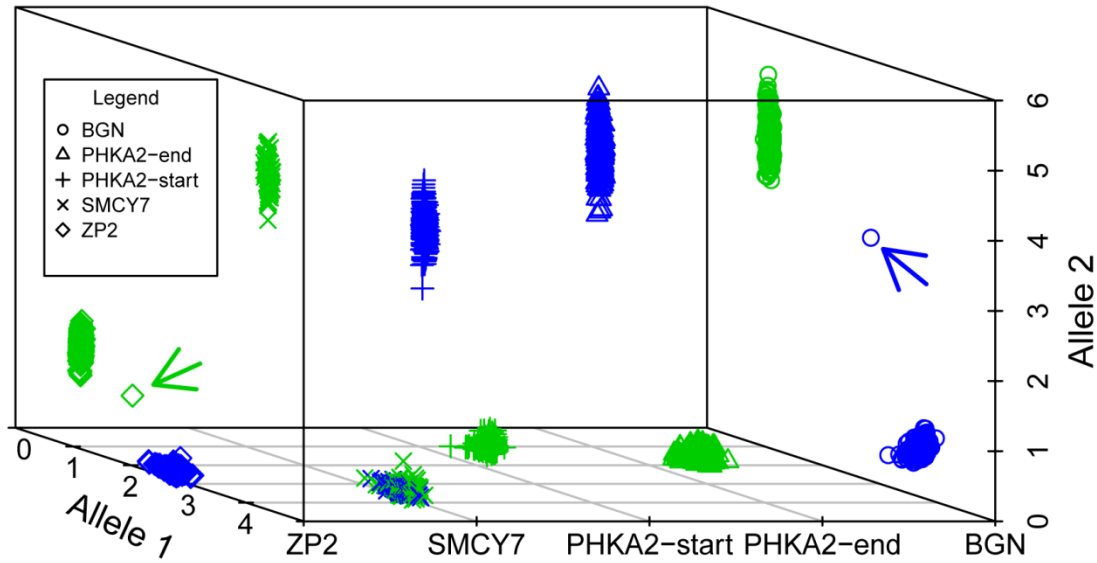
Supplementary Figure S2 | Genetic differentiation of Thai *versus* Myanmar individuals based on five nuclear SNPs. Genotyping of 5 SNPs from 196 Myanmar (blue) and 463 Thai (green) *Craseonycteris thonglongyai* individuals (ZP2 is an autosomal marker; SMCY7 is a Y-Chromosome marker and PHKA2-start, PHKA2-end and BGN are X-Chromosome markers). Note the presence of two heterozygote individuals (arrows), one for ZP2 and one for BGN. For SMCY7, the two alleles are present in the Thai population.
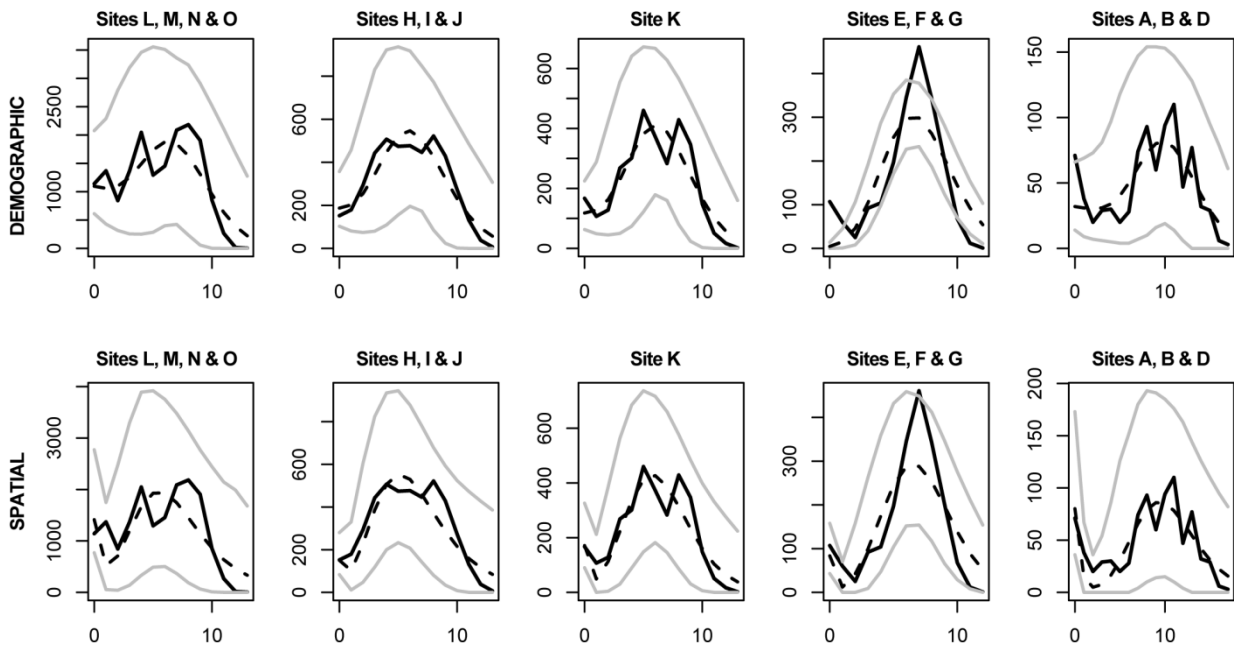
Supplementary Figure S3 | Observed and predicted mismatch distributions under a model of demographic expansion (top), or spatial expansion (bottom). For the five groups of sites, a model of constant population size was rejected as the 99% confidence intervals of θ_0 and θ_1 were not overlapping. Solid black lines represent the observed frequency of pairwise distributions, dashed black lines indicate the expected distribution under the model and the solid grey lines indicate the upper and lower 95% confidence interval for the expected. For the demographic expansion (top panels), the raggedness index (Rag.) and associated P-values are as follow, sites L, M, N & O: Rag.= 0.0129, p=0.603; sites H, I & J: Rag.= 0.006, p=0.943; site K: Rag.=0.0151, p=0.378; sites E, F & G: Rag.=0.0388, p=0.002; sites A, B & D: Rag.=0.019, p=0.246. For the spatial expansion (bottom panels), the raggedness index and associated P-values are as follows, sites L, M, N & O: Rag.= 0.0129, p=0.881; sites H, I & J: Rag.= 0.006, p=0.96; site K: Rag.=0.0151, p=0.747; sites E, F & G: Rag.=0.0388, p=0.17; sites A, B & D: Rag.=0.019, p=0.857.
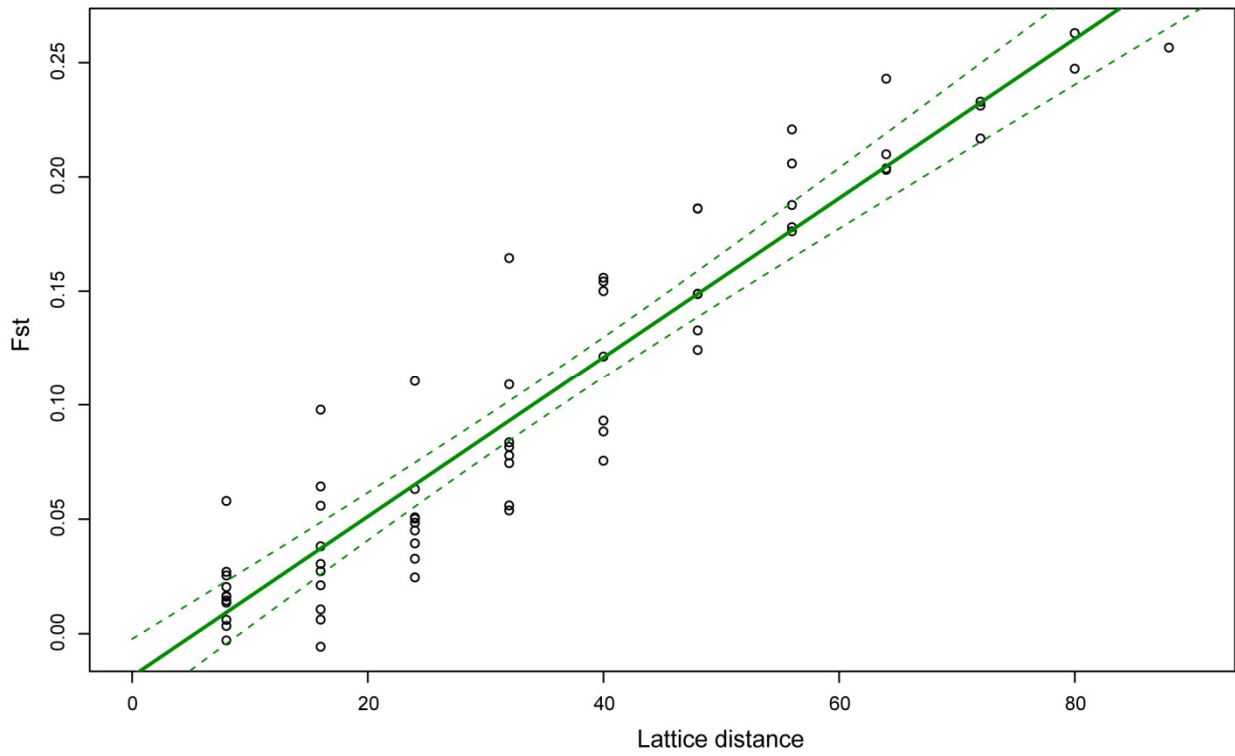
Supplementary Figure S4 | Example of isolation by distance obtained by simulating a spatially expanding population. Simulations were carried out in SPLATCHE with the migration rate and growth parameter values set to 0.3 each and the carrying capacity set to 250 (Mantel test; r=0.95, R^2=0.90, P<0.001). The program was run over 400 generations (see Supplementary Table S3).
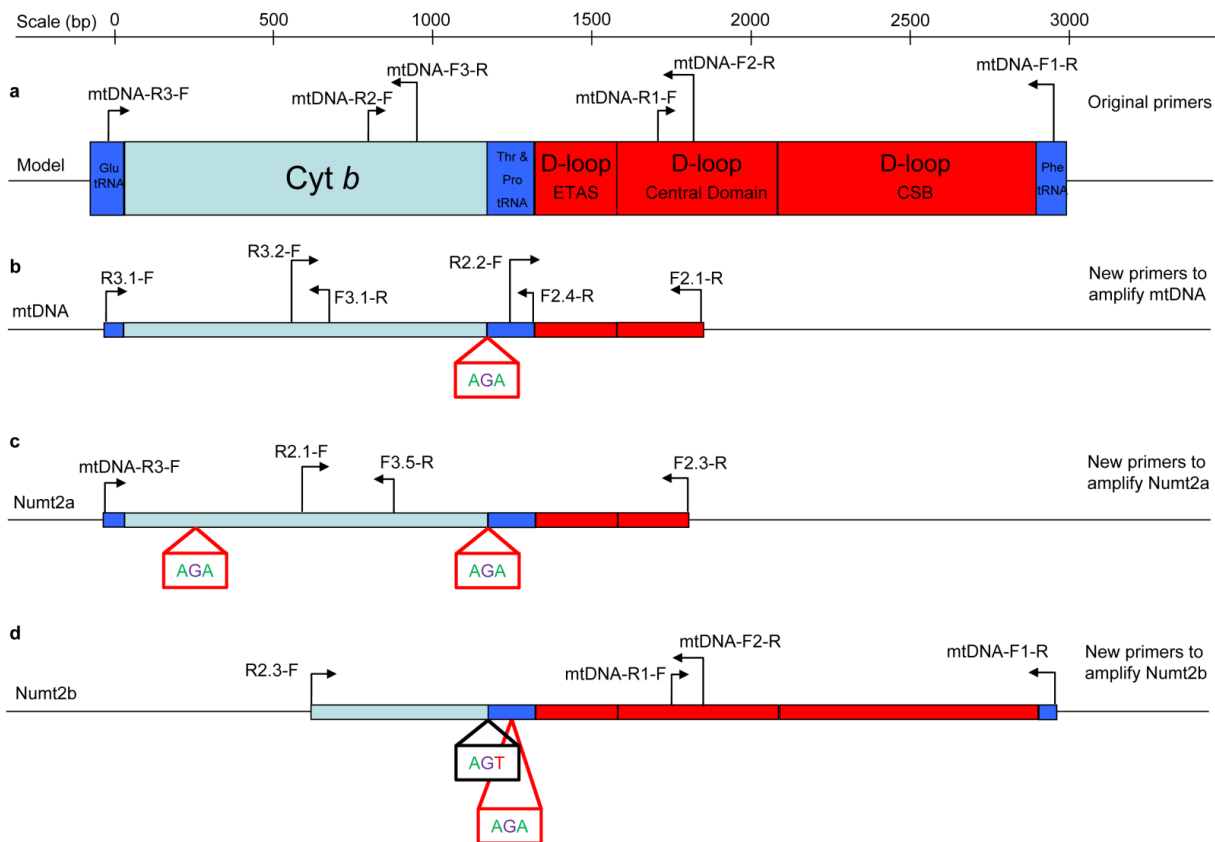
Supplementary Figure S5 | Schematic diagram of the organisation of the mitochondrial DNA in mammals. The D-loop domains were specified according to Sbisà *et al.*[61] although the relative length of each domain can vary between species or individuals of the same species. Each arrow indicates a primer with its name specified. Red boxes indicate the presence of a stop codon (AGA; mtDNA code) and black boxes, the absence of a stop codon where one would be expected (see text for further explanations).

Supplementary Table S1 | List of *Craseonycteris thonglongyai* haplotypes from nuclear markers and their corresponding GenBank accession numbers.

Individual reference	Country	Haplotype	Accession N.	Origin	Marker
CT1	Myanmar	Cratho_Numt2a_1	GU247603	nuclear	Num2a
CT255	Thailand	Cratho_Numt2a_2	GU247604	nuclear	Num2a
CT17	Myanmar	Cratho_Numt2b_1	GU247605	nuclear	Numt2b
CT255	Thailand	Cratho_Numt2b_2	GU247606	nuclear	Numt2b
CT18	Myanmar	Cratho_ZP2_1	GU247611	nuclear	ZP2
CT256	Thailand	Cratho_ZP2_2	GU247612	nuclear	ZP2
CT18	Myanmar	Cratho_BGN_1	GU247601	X-chromosome	BGN
CT803	Thailand	Cratho_BGN_2	GU247602	X-chromosome	BGN
CT17	Myanmar	Cratho_PHKA2_1	GU247607	X-chromosome	PHKA2
CT255	Thailand	Cratho_PHKA2_2	GU247608	X-chromosome	PHKA2
CT18	Myanmar	Cratho_SMCY_1	GU247609	Y-chromosome	SMCY7
CT255	Thailand	Cratho_SMCY_2	GU247610	Y-chromosome	SMCY7

Supplementary Table S2 | List of *Craseonycteris thonglongyai* mitochondrial DNA haplotypes (*Cytb*, tRNA-Pro, tRNA-Thr, D-loop) and their corresponding GenBank accession numbers.

Individual reference	Country	Haplotype name	Frequency	Accession number
CT13	Myanmar	Cratho_1	17	GU247613
CT16	Myanmar	Cratho_2	10	GU247683
CT17	Myanmar	Cratho_3	10	GU247614
CT28	Myanmar	Cratho_4	89	GU247684
CT36	Myanmar	Cratho_5	1	GU247615
CT38	Myanmar	Cratho_6	2	GU247685
CT54	Myanmar	Cratho_7	1	GU247616
CT60	Myanmar	Cratho_8	1	GU247686
CT88	Myanmar	Cratho_9	1	GU247617
CT89	Myanmar	Cratho_10	2	GU247687
CT90	Myanmar	Cratho_11	1	GU247618
CT111	Myanmar	Cratho_12	1	GU247688
CT114	Myanmar	Cratho_13	1	GU247619
CT118	Myanmar	Cratho_14	1	GU247689
CT298	Myanmar	Cratho_15	1	GU247620
CT229	Thailand	Cratho_16	13	GU247690
CT230	Thailand	Cratho_17	40	GU247621
CT232	Thailand	Cratho_18	22	GU247691
CT233	Thailand	Cratho_19	6	GU247622
CT235	Thailand	Cratho_20	7	GU247692
CT236	Thailand	Cratho_21	4	GU247623
CT237	Thailand	Cratho_22	1	GU247693
CT238	Thailand	Cratho_23	13	GU247624
CT239	Thailand	Cratho_24	3	GU247694
CT240	Thailand	Cratho_25	2	GU247625
CT242	Thailand	Cratho_26	4	GU247695
CT243	Thailand	Cratho_27	3	GU247626
CT245	Thailand	Cratho_28	1	GU247696
CT255	Thailand	Cratho_29	8	GU247627
CT256	Thailand	Cratho_30	7	GU247697
CT265	Thailand	Cratho_31	1	GU247628
CT266	Thailand	Cratho_32	3	GU247698

Individual reference	Country	Haplotype name	Frequency	Accession number
CT267	Thailand	Cratho_33	6	GU247629
CT377	Thailand	Cratho_34	11	GU247699
CT378	Thailand	Cratho_35	8	GU247630
CT379	Thailand	Cratho_36	4	GU247700
CT382	Thailand	Cratho_37	8	GU247631
CT383	Thailand	Cratho_38	21	GU247701
CT388	Thailand	Cratho_39	1	GU247632
CT390	Thailand	Cratho_40	5	GU247702
CT392	Thailand	Cratho_41	13	GU247633
CT393	Thailand	Cratho_42	1	GU247703
CT394	Thailand	Cratho_43	6	GU247634
CT395	Thailand	Cratho_44	2	GU247704
CT397	Thailand	Cratho_45	12	GU247635
CT398	Thailand	Cratho_46	3	GU247705
CT399	Thailand	Cratho_47	1	GU247636
CT400	Thailand	Cratho_48	7	GU247706
CT405	Thailand	Cratho_49	2	GU247637
CT406	Thailand	Cratho_50	1	GU247707
CT408	Thailand	Cratho_51	3	GU247638
CT410	Thailand	Cratho_52	4	GU247708
CT411	Thailand	Cratho_53	1	GU247639
CT413	Thailand	Cratho_54	7	GU247709
CT414	Thailand	Cratho_55	5	GU247640
CT415	Thailand	Cratho_56	7	GU247710
CT418	Thailand	Cratho_57	2	GU247641
CT423	Thailand	Cratho_58	1	GU247711
CT425	Thailand	Cratho_59	6	GU247642
CT435	Thailand	Cratho_60	2	GU247712
CT437	Thailand	Cratho_61	3	GU247643
CT438	Thailand	Cratho_62	3	GU247713
CT443	Thailand	Cratho_63	1	GU247644
CT445	Thailand	Cratho_64	2	GU247714
CT447	Thailand	Cratho_65	1	GU247645
CT448	Thailand	Cratho_66	2	GU247715
CT450	Thailand	Cratho_67	1	GU247646
CT453	Thailand	Cratho_68	1	GU247716

Individual reference	Country	Haplotype name	Frequency	Accession number
CT454	Thailand	Cratho_69	1	GU247647
CT455	Thailand	Cratho_70	2	GU247717
CT456	Thailand	Cratho_71	1	GU247648
CT459	Thailand	Cratho_72	2	GU247718
CT463	Thailand	Cratho_73	1	GU247649
CT466	Thailand	Cratho_74	8	GU247719
CT471	Thailand	Cratho_75	1	GU247650
CT475	Thailand	Cratho_76	1	GU247720
CT478	Thailand	Cratho_77	7	GU247651
CT482	Thailand	Cratho_78	2	GU247721
CT488	Thailand	Cratho_79	1	GU247652
CT493	Thailand	Cratho_80	2	GU247722
CT496	Thailand	Cratho_81	1	GU247653
CT498	Thailand	Cratho_82	1	GU247723
CT500	Thailand	Cratho_83	3	GU247654
CT510	Thailand	Cratho_84	1	GU247724
CT515	Thailand	Cratho_85	2	GU247655
CT516	Thailand	Cratho_86	10	GU247725
CT517	Thailand	Cratho_87	6	GU247656
CT518	Thailand	Cratho_88	11	GU247726
CT523	Thailand	Cratho_89	6	GU247657
CT524	Thailand	Cratho_90	3	GU247727
CT527	Thailand	Cratho_91	5	GU247658
CT529	Thailand	Cratho_92	1	GU247728
CT531	Thailand	Cratho_93	3	GU247659
CT532	Thailand	Cratho_94	1	GU247729
CT534	Thailand	Cratho_95	3	GU247660
CT539	Thailand	Cratho_96	10	GU247730
CT541	Thailand	Cratho_97	3	GU247661
CT552	Thailand	Cratho_98	1	GU247731
CT554	Thailand	Cratho_99	1	GU247662
CT555	Thailand	Cratho_100	2	GU247732
CT580	Thailand	Cratho_101	1	GU247663
CT592	Thailand	Cratho_102	1	GU247733
CT595	Thailand	Cratho_103	2	GU247664
CT596	Thailand	Cratho_104	1	GU247734

Individual reference	Country	Haplotype name	Frequency	Accession number
CT608	Thailand	Cratho_105	1	GU247665
CT611	Thailand	Cratho_106	1	GU247735
CT614	Thailand	Cratho_107	1	GU247666
CT617	Thailand	Cratho_108	1	GU247736
CT623	Thailand	Cratho_109	1	GU247667
CT628	Thailand	Cratho_110	6	GU247737
CT635	Thailand	Cratho_111	1	GU247668
CT653	Thailand	Cratho_112	1	GU247738
CT670	Thailand	Cratho_113	2	GU247669
CT671	Thailand	Cratho_114	2	GU247739
CT672	Thailand	Cratho_115	1	GU247670
CT674	Thailand	Cratho_116	1	GU247740
CT675	Thailand	Cratho_117	1	GU247671
CT676	Thailand	Cratho_118	1	GU247741
CT681	Thailand	Cratho_119	1	GU247672
CT682	Thailand	Cratho_120	1	GU247742
CT686	Thailand	Cratho_121	3	GU247673
CT688	Thailand	Cratho_122	1	GU247743
CT692	Thailand	Cratho_123	3	GU247674
CT700	Thailand	Cratho_124	2	GU247744
CT726	Thailand	Cratho_125	1	GU247675
CT735	Thailand	Cratho_126	2	GU247745
CT737	Thailand	Cratho_127	3	GU247676
CT747	Thailand	Cratho_128	2	GU247746
CT763	Thailand	Cratho_129	1	GU247677
CT779	Thailand	Cratho_130	1	GU247747
CT780	Thailand	Cratho_131	1	GU247678
CT785	Thailand	Cratho_132	1	GU247748
CT792	Thailand	Cratho_133	1	GU247679
CT793	Thailand	Cratho_134	1	GU247749
CT794	Thailand	Cratho_135	3	GU247680
CT796	Thailand	Cratho_136	1	GU247750
CT799	Thailand	Cratho_137	1	GU247681
CT800	Thailand	Cratho_138	1	GU247751
CT801	Thailand	Cratho_139	1	GU247682

Supplementary Table S3 | Results of the spatial expansion simulations run in SPLATCHE for different parameters (C, m and r) combinations. A total of 24 simulations were run with 100 replicates each. The 'Generations' column specify how many generations were necessary to fill in the lattice. The 'Match observed' column present the simulations in which an isolation by distance pattern similar to the one seen in *Craseonycteris* was observed. See simulation details in the "simulation of a spatially expanding population" section of the Supplementary Material.

C	m	r	Generations	Match observed
250	0.1	0.1	1100	Yes
250	0.1	0.3	600	Yes
250	0.1	0.5	500	No
250	0.3	0.1	700	Yes
250	0.3	0.3	400	Yes
250	0.3	0.5	350	Yes
250	0.5	0.1	550	Yes
250	0.5	0.3	350	Yes
250	0.5	0.5	300	Yes
250	0.7	0.1	500	No
250	0.7	0.3	300	Yes
250	0.7	0.5	250	Yes
100	0.1	0.1	1150	Yes
100	0.1	0.3	600	No
100	0.1	0.5	500	No
100	0.3	0.1	700	Yes
100	0.3	0.3	400	Yes
100	0.3	0.5	350	Yes
100	0.5	0.1	550	Yes
100	0.5	0.3	350	Yes
100	0.5	0.5	300	Yes
100	0.7	0.1	500	No
100	0.7	0.3	300	Yes
100	0.7	0.5	250	Yes

Supplementary Table S4 | Evaluation of the seven organisational models considering geographic distance, echolocation distance, the presence of barriers and genetic distance.
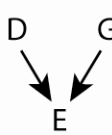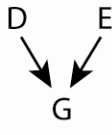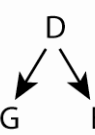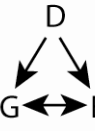
Model and expectation[a]	Mantel r	P-value[b]	Expectation fit
Model 1, effect of distance			
DG.B $\neq$ 0	0.95	<0.001*	Yes
DG.E $\neq$ 0	0.94	<0.001*	Yes
BG.D = 0	0.29	0.023 NS	Yes
EG.D = 0	0.43	0.003*	No
Model 2, effect of barrier			
BG.D $\neq$ 0	0.29	0.023 NS	No
BG.E $\neq$ 0	0.00	0.884 NS	No
DG.B = 0	0.95	<0.001*	No
EG.B = 0	0.44	0.003*	No
Model 3, effect of echolocation			
EG.B $\neq$ 0	0.44	0.003*	Yes
EG.D $\neq$ 0	0.43	0.003*	Yes
BG.E = 0	0.00	0.884 NS	Yes
DG.E = 0	0.94	<0.001*	No
Model 4, effect of distance and barrier			
BG.E $\neq$ 0	0.00	0.884 NS	No
DG.E $\neq$ 0	0.94	<0.001*	Yes
BG.D $\neq$ 0	0.29	0.023 NS	No
DG.B $\neq$ 0	0.95	<0.001*	Yes
EG.B = 0	0.44	0.003*	No
EG.D = 0	0.43	0.003*	No
Model 5, effect of distance and echolocation			
EG.B $\neq$ 0	**0.44**	**0.003***	**Yes**
DG.B $\neq$ 0	**0.95**	**<0.001***	**Yes**
EG.D $\neq$ 0	**0.43**	**0.003***	**Yes**
DG.E $\neq$ 0	**0.94**	**<0.001***	**Yes**
BG.E = 0	**0.00**	**0.884 NS**	**Yes**
BG.D = 0	**0.29**	**0.023 NS**	**Yes**
Model 6, effect of barrier and echolocation			
BG.D $\neq$ 0	0.29	0.023 NS	No
EG.D $\neq$ 0	0.43	0.003*	Yes
BG.E $\neq$ 0	0.00	0.884 NS	No
EG.B $\neq$ 0	0.44	0.003*	Yes
DG.B = 0	0.95	<0.001*	No
DG.E = 0	0.94	<0.001*	No
Model 7, effect of distance, barrier and			
BG.D $\neq$ 0	0.29	0.023 NS	No
BG.E $\neq$ 0	0.00	0.884 NS	No
DG.B $\neq$ 0	0.95	<0.001*	Yes
DG.E $\neq$ 0	0.94	<0.001*	Yes
EG.B $\neq$ 0	0.44	0.003*	Yes
EG.D $\neq$ 0	0.43	0.003*	Yes

Boldface indicates fully supported models,

[a] B = barrier, G = genetics, D = distance, E = echolocation. The period in the expectation abbreviations separates the covariate matrix from the two primary matrices (i.e. DG.B indicates a Mantel test between the distance and genetic matrices, with the barrier matrix partialed out),

[b] *P-values* are preceded with the sign '*' if significant or 'NS' if non-significant after sequential Bonferroni correction.

Supplementary Table S5 | Evaluation of the six models of causal relationships between geographic distance (D), echolocation distance (E) and genetic distance (G).

Model and expectation	Expectation fit	Model and expectation	Expectation fit
Model 1		**Model 4**	
DE ≠ 0	**Yes**	**DG ≠ 0**	**Yes**
EG ≠ 0	**Yes**	**EG ≠ 0**	**Yes**
DG = 0	No	**DG ≥ DE**	**Yes**
DG.E ≠ 0	**Yes**	**DE.G = 0**	**Yes**
EG.D ≠ 0	**Yes**	**DG.E ≠ 0**	**Yes**
DE.G ≠ 0	No	**EG.D ≠ 0**	**Yes**
DE.G ≥ DE	No	**DE.G ≤ DG**	**Yes**
EG.D ≥ EG	No	**EG.D ≤ ED**	**Yes**
Model 2		**DG x EG ≈ DE**	**Yes**
DG ≠ 0	**Yes**	Model 5	
EG ≠ 0	**Yes**	DE ≠ 0	**Yes**
DE = 0	No	DG ≠ 0	**Yes**
DE.G ≠ 0	No	DE.G ≠ 0	No
EG.D ≠ 0	**Yes**	DG.E ≠ 0	**Yes**
DG.E ≠ 0	**Yes**	EG.D = 0	No
DG.E ≥ DG	**Yes**	DG.E ≤ DG	**Yes**
EG.D ≥ EG	No	DE.G ≤ DE	**Yes**
Model 3		DE x DG ≈ EG	**Yes**
DE ≠ 0	**Yes**	Model 6	
EG ≠ 0	**Yes**	DE ≠ 0	**Yes**
DE ≥ DG	No	DG ≠ 0	**Yes**
DG.E = 0	No	EG ≠ 0	**Yes**
DE.G ≠ 0	No	DE.G ≠ 0	No
EG.D ≠ 0	**Yes**	DG.E ≠ 0	**Yes**
DE.G ≤ DE	**Yes**	EG.D ≠ 0	**Yes**
EG.D ≤ DG	**Yes**		
DE x EG ≈ DG	No		

Boldface indicates fully supported models.

Abbreviations are identical to those used in Supplementary Table S4.

Supplementary Table S6 | Summary of the synteny between the *Homo sapiens* block including the *RBP-J* gene (left side) and other species (right side) available in Ensembl (http://www.ensembl.org/index.html). The microsatellite CTC1 blasts a region 50 kb at the 5'-side of *RBP-J*, that is at position 26.267 in Chromosome 4 of *Homo sapiens* (GRCh37 primary reference assembly). In all species looked at, the synteny around *RBP-J* was conserved. Base position or block length are expressed in Mega-bases (1M=1,000,000 bases).

Reference species (Chromosome No.)	Syntenic block (M)			Target species (Chromosome No.)	Syntenic block (M)		
	Start	End	Size		Start	End	Size
Homo sapiens (4)	23.8	27.4	3.6	*Gallus gallus (4)*	75.2	76.7	1.5
Homo sapiens (4)	0	44.8	44.8	*Pan troglodytes (4)*	0	45.3	45.3
Homo sapiens (4)	17.5	49.1	31.6	*Bos taurus (6)*	37.9	70.3	32.4
Homo sapiens (4)	17.8	41.3	23.5	*Canis familiaris (3)*	74.2	94.2	20
Homo sapiens (4)	9.8	49.1	39.3	*Equus caballus (3)*	80	112.5	32.5
Homo sapiens (4)	9.8	49.1	39.3	*Macaca mulatta (5)*	4.6	44.4	39.8
Homo sapiens (4)	4.2	49.1	44.9	*Mus musculus (5)*	35.7	73.8	38.1
Homo sapiens (4)	10	48.9	38.9	*Monodelphis domestica (5)*	174.2	220.4	46.2
Homo sapiens (4)	9.8	49.3	39.5	*Pongo pygmaeus (4)*	9.2	50.7	41.5
Homo sapiens (4)	10.8	49.1	38.3	*Sus scrofa (8)*	4.6	33.5	28.9
Homo sapiens (4)	17.5	28.3	10.8	*Ornithorhynchus anatinus (18)*	1.1	6.6	5.5
Homo sapiens (4)	26.1	33.7	7.6	*Rattus norvegicus (14)*	53.6	62.3	8.7

Supplementary Table S7 | Complete mitochondrial DNA sequences of nine bat species downloaded from GenBank.

Seq. N°	Genus	Species	gi	Author/Reference
NC_007393	*Rousettus*	*aegyptiacus*	74310519	Omatsu *et al.,* unpublished
NC_006925	*Mystacina*	*tuberculata*	62184382	NCBI Genome Project, unpublished
NC_002612	*Pteropus*	*dasymallus*	11386118	Nikaido *et al.,* unpublished
NC_005434	*Rhinolophus*	*pumilus*	42632271	[62]
NC_005436	*Pipistrellus*	*abramus*	42632257	[62]
NC_002626	*Chalinolobus*	*tuberculatus*	11610804	[63]
NC_002619	*Pteropus*	*scapulata*	11602891	[63]
NC_005433	*Rhinolophus*	*monoceros*	42717961	[64]
NC_002009	*Artibeus*	*jamaicensis*	5835666	[65]

Supplementary Table S8 | Primer names and sequences used in the present study.

Primer name	Primer sequence (5' - 3')	Reference
mtDNA-R3-F	TGGCATGAAAAATCACCGTTGT	This study
mtDNA-F3-R	AGGATGGCGTATGCAAATAGGAA	This study
mtDNA-R2-F	CCGACCTATTAGGAGACCCAGA	This study
mtDNA-F2-R	ATGGCCCTGAAGAAAGAACCAGATG	This study
mtDNA-R1-F	CTCACGTGAAACCAGCAACC	This study
mtDNA-F1-R	ATACTCATCTAGGCATTTTCAGTGC	This study
R3.1-F	TGAAAAACCATCGTTGTATTTCAACTACAA	This study
F3.1-R	CGGTTGGGTTATTGGACCCA	This study
R3.2-F	AGAATGAGTCTGAGGTGGCTTTT	This study
F2.4-R	TTCCTTGAAGTCTTTGGAGAATG	This study
R2.2-F	CCTAGTTCTTATACCCCTAGCAGGA	This study
F2.1-R	GACACATGGTTCAAGTTAAGCTCAG	This study
BGN-F	CTCCAAGAACCACCTGGTG	[66]
BGN-R	TTCAAAGCCACTGTTCTCCAG	[66]
PHKA2-F	GTGGGAGCGTGGAGATAAGA	[67]
PHKA2-R	TGAATCACTGACTTGCGTCC	[67]
SMCY7-F	TGGAGGTGCCCRAARTGTA	[68]
SMCY7-R	AACTCTGCAAASTRTACTCCT	[68]
ZP2-E12F	GACGGATCTTCCTCCAAGC	This study
ZP2-E14R	CATCCACGACAATGTTCCAC	This study

Supplementary Table S9 | Forward and reverse primer and Taqman probe sequences for five nuclear introns SNPs. SNPs were fluorescently labelled using either FAM or VIC dyes.

SNP	Primer direction	Primer 5' - 3'	Probe dye	Probe sequence
PHKA 2-start	Fwd	CACCTACAGGTCTATGCTTTTGCT	VIC	CTGGATGTATGATATCCA
	Rev.	CCTTACCCTACAAAGTGGTCATATAGTT	FAM	CTGGATGTATGATGTCCA
PHKA 2-end	Fwd	AGCAGCCTATAAGAAAGCTCTAATCTACT	VIC	CTAACCAGTCATATGGTC
	Rev.	AAAGGCGATGAAACACAATGCTAAA	FAM	ACCAGTCGTATGGTC
BGN	Fwd	GGTCTTGAACAGCTAGGAGTTAGTG	VIC	CCCCCGAAGCAGTT
	Rev.	GCACCTGCTCCCCTATTCG	FAM	CCCCGGAGCAGTT
SMCY 7	Fwd	TGATAAGAATGGGTGGACAGAAGAGT	VIC	TCTTGCTAAAGGTTCTAC
	Rev.	CCCTCAGATTTCTGATCCTATAGTTTATTTCAT	FAM	TTGCTAACGGTTCTAC
ZP2	Fwd	GCCTGCCCTCCAATGAACT	VIC	CTGAAGCTGCCAGGCT
	Rev.	GGGCAACAGGAGCTTAAGCAT	FAM	TGAAGCTGCCGGGCT

Supplementary Note 1

Echolocation and mate choice. Our data show that echolocation does have a minimum effect on gene flow between two populations in close vicinity (Northern/Central *versus* Southern populations) but how can this happen? Most bat species predominately rely on echolocation for sensory perception[22]. Throughout the historical literature, echolocation has been perceived to have a dominant role in orientation and prey capture. A bat's larynx, pinnae shape and inner ear structures are directly correlated with their echolocation capacities and therefore foraging capabilities[22]. However, not much is known about how echolocation can influence social communication in bats (however, see[14,115]). Recent literature has shown that bats can discriminate individuals on the basis of their echolocation call[12,116-118] and some studies have suggested that echolocation parameters somehow influence mate choice preference in bats[5,119,120]. This is not that surprising. Indeed, it has been shown that bird beaks are predominantly linked to their foraging capabilities but also play a role in determining song structure which then in turn affects their mating songs[121], and thus mate choice. Therefore, in bats that also use sound for mating rituals and courtship[122], it seems likely that the apparatus they use to make these sounds (larynx) and to perceive these sounds (their outer and inner ears), which are directly linked to their echolocation capacities, will play a role in determining their mate preferences. It has been shown that in *Saccopteryx bilineata*, male reproductive success as determined by their number of offspring was directly correlated with echolocation frequency parameters[123]. Further studies on *Myotis lucifugus* have demonstrated that males preferred calls of females who mated frequently rather than calls of females rarely mating[124]. The European cryptic pipistrelles (*Pipistrellus pipistrellus* and *P. pygmaeus*) echolocate with an average difference of 10 kHz[120] and also show frequency differences in their social calls[125]. Therefore echolocation can indeed play a role in communication and mate choice and our data suggest that bats are choosing to mate with other bats that have the same 'echolocation' call type.

Supplementary Note 2

Echolocation competition. Echolocating bats produce ultrasonic signals and determine the direction, distance, and features of objects in the environment from the arrival time, amplitude, and spectrum of sonar reflections[13]. This process commonly named 'echolocation' is a very challenging task requiring specific adaptations of the vocal and auditory system as well as the brain[45,107,108]. We present below two hypotheses (Resource partitioning and Interference) potentially explaining the observed pattern of echolocation call variation observed in *C. thonglongyai* in Thailand. It is important to note that these two hypotheses are not mutually exclusive and could act together to strengthen the divergent selection.

(A) **Resource Partitioning:** Assuming targets are spheres and a the speed of sound equals 347.65 m s-1 (for a temperature of 25°C and a relative humidity of 80%), the theory predicts that, *M. siligorensis* echolocating at 70 kHz can detect targets of 4.96 mm whereas *C. thonglongyai* echolocating at 76 kHz can detect targets of smaller size (4.56 mm). This would mean that *C. thonglongyai* can detect preys that are 8% smaller than the smallest prey detectable by *M. siligorensis*. Although this does not correspond to a large difference, not enough is currently known about the complexities of bat echolocation and their perception to rule out the effect of this 6 kHz difference on resource partitioning (although, see[43]). Furthermore, experimental studies broadcasting ultrasounds at real insects and recording returning echoes have showed that these calculations were not very reliable as insects are not spheres[109]. The strength of the returning echo does not only depend on the insect size but also on its wing beat and the angle (insect ensonified from the front, side or back)[109]. Therefore it is possible that this shift in call is biologically meaningful but we do not know enough to accept or reject this hypothesis yet[41].

(B) **Interference:** Bats are also faced with the challenge of separating their calls from the potential interference of various sounds in the environment where they are navigating. In echolocating bats, the onset of the emitted call activates a gating mechanism that establishes a time window during which pulse-echo pairs are processed for target distance determination[45]. This process has been validated in bats with different echolocation types[44,45,110] and should apply to all echolocating bat species[45]. Experiments by Roverud & Grinnel[44] on *Noctilio albiventris*

20

demonstrated that playing artificial pulses resembling the bat's calls interfered with the bat's ability to determine the distance of objects. The disrupting effect of the artificial sound is likely due to the interference with the bat's processing of information from its own sounds by stimulating the same population of neurons that extract distance information from the bat's echolocation sound[22,110]. Interference occurred only when the constant frequency of the playback sound was between 2 to 2.5 kHz above and 5 kHz below the frequency of the beginning of the frequency modulated sweep[44], demonstrating a narrow frequency window for interference. Assuming these thresholds are similar for our study species implies that *C. thonglongyai* echolocating at 75 kHz or above (Southern populations) avoids interference from *Myotis siligorensis* present in the area and echolocating at 70 kHz. In the Northern populations, where *M. siligorensis* has not been found, *C. thonglongyai* echolocates at 73-74 kHz, which is within the interference frequency window.

Intra-specific jamming in *C. thonglongyai* could be avoided by temporal shifts in frequency as demonstrated in many other bat species. When two or more bats of the same species are flying within the same airspace (within earshot of one another) individuals adjust the frequencies dominating their echolocation calls to avoid jamming each other sonar[111-114]. Nevertheless, Duanghkae[34] showed that in *Craseonycteris*, each individual had its own foraging area and occasionally, when another bat got into an individual's foraging area, it would be chased out. This suggests that *Craseonycteris* might avoid intra-specific jamming by avoiding contact with other individuals in the foraging grounds.

Supplementary Methods

Distinction of mtDNA *versus* NUMTs. Complete mitochondrial DNA sequences of nine bat species from five families were downloaded from GenBank (Supplementary Table S7) and aligned using MEGA version 3.1[69]. The D-loop was visually checked and realigned by hand. Six conserved primers were designed to amplify the entire *Cytb* and the D-loop in three overlapping fragments (see Supplementary Fig. S5 for primer locations and Supplementary Table S8 for primer sequences). Reactions were carried out in 25 μL simplex reactions containing 2 μL of DNA extract (at 2-5 ng/μL), 1X PCR buffer minus Mg (Invitrogen), 1.5 mM MgCl$_2$, 0.4 μM each primer, 0.2 mM dNTPs and 1 U Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen). PCR volumes and reagents above-mentioned were used for all PCRs unless otherwise stated.

To ascertain that we amplified the mitochondrial DNA and not Numts (nuclear copies of mitochondrial DNA), PCRs were completed with all possible primer combinations to ensure that all primers bound only to the mitochondrial DNA. All PCRs were carried out using the DNA from the same individual to avoid potential inter-individual differences. Identical PCR cycling conditions were used for all primer combinations; initial step 10' at 95°C, then 10 cycles of 15" at 95°C, 30" at 60°C (reduce by 2°C every 2 cycles), 1' at 72°C, following by 30 cycles of 15" at 95°C, 30" at 50°C and 1' at 72°C and a final step for 10' at 72°C.

Using the DNA from a single individual and different combinations of primers, three different sequences were obtained for the same targeted mtDNA region. The PCRs, repeated on seven different individuals, furnished consistent results, suggesting the presence of Numts[70,71]. The three different sequences were aligned and new primers were designed to specifically amplify each fragment (see primers represented in Supplementary Fig. S5 and Supplementary Table S8). Two amplified fragments spanned from the tRNA-Glu to the Central Domain of the D-loop as targeted whereas for one fragment, the first 730 bp of the *Cytb* could not be amplified (Supplementary Fig. S5d). When translating the amino-acid sequence of the *Cytb* using the mammalian mitochondrial genetic code, one stop codon was found in one sequence thus classified as Numt (Numt2a) (Supplementary Fig. S5c). A second sequence, for which we were not able to amplify the beginning of *Cytb*, presented a 2 bp deletion at position 1140-

1141 when compared to the other two sequences. This deletion created a frame-shift modifying the *Cytb* stop codon which was then found in position 1183-1185 (Supplementary Fig. S5d). Therefore, the tRNA-Thr present on the 3' side of the *Cytb* was reduced by 45 bp. This second sequence was thus also considered as a Numt (Numt2b). This was further confirmed by a simple sequence divergence table whereby Numt sequences from Thailand and Myanmar individuals were quasi-identical (<0.1% divergence) whereas true mitochondrial DNA from Thailand and Myanmar were more different (>1% divergence). This difference between mitochondrial DNA and Numts reflects the expected higher mutation rate of mtDNA *versus* nuclear DNA[72,73].

Mitochondrial DNA amplification and sequencing. A 1840 bp mitochondrial DNA fragment encompassing the entire *Cytb*, tRNA Threonine, tRNA Proline and part of the D-loop was amplified by PCR in three overlapping fragments using three primer pairs R3.1-F/F3.1-R, R3.2-F/F2.4-R and R2.2-F/F2.1-R (Supplementary Table S8 and Supplementary Fig. S5b) and the PCR recipe described above. Amplifications were carried out in a DNA Engine DYAD^{TM} thermocycler (MJ Research) with the following PCR program: initial step 10' at 95°C, then 10 cycles of 15" at 95°C, 30" at 65°C (reduce by 2°C every 2 cycles), 1' at 72°C, following by 30 cycles of 15" at 95°C, 30" at 55°C and 1' at 72°C and a final step for 10' at 72°C. The amplified products were then purified and directly sequenced by Macrogen (Korea) using the same primers as mentioned above. The six sequences per sample were edited and assembled using the program Sequencher 4.7 (Gene Codes Corporation, Ann Arbor, MI).

All the 468 samples from Thailand were sequenced for the entire mitochondrial DNA fragment. In Myanmar, all samples were sequenced except from one colony (M13) where only 65 of the 128 samples were used. Out of 65 samples sequenced in M13, only 8 unique haplotypes were found, suggesting that we already detected all unique haplotypes[74]. Individuals sampled twice (see the 'Microsatellites genotyping and analysis' paragraph in the Methods) were removed, leaving a data set comprising 463 sequences for Thailand and 139 for Myanmar.

Phylogenetic reconstruction and dating. Using mitochondrial, Numt2a and Numt2b sequence, phylogenetic reconstruction was undertaken using the Bayesian inference in BEAST[54]. The part of the D-

loop containing insertions-deletions was removed from the alignment prior to phylogenetic reconstruction due to difficulty in aligning this section, leaving a 1.8 kb alignment. The general time-reversible + gamma-distributed rates among sites + proportion of invariant sites (GTR+ Γ+I) substitution model was used as determined by ModelTest version 3.7[75]. A strict molecular clock model was preferred over a relaxed molecular clock model as advised by Drummond *et al.*[76] when the standard deviation of the uncorrelated lognormal relaxed clock (parameter ucld.stdev in BEAST) is smaller than 1, which is what was observed when analysing the present dataset. Therefore, a strict molecular clock model was applied with a fixed mean substitution rate of 2.21×10^{-8} subs/site/year. The mutation rate over the entire sequence was calculated by weighting the rate contribution of the various functional regions Cytb, tRNAs, ETAS and Central domain[77] by their relative length. ETAS and Central domain were defined according to the alignments provided by Sbisà *et al.*[61] whereas tRNAs were delimited using the online program ARWEN version 1.2[78]. The average mutation rate was estimated to be 2.21×10^{-8} subs/site/year, which equals to a divergence rate of 4.42%/Myr. No outgroup was specified and the constant size coalescent was used as a tree prior. The program was run for 40,000,000 generations and sampled every 500. The first 4,000,000 generations were discarded as burn-in. Effective sample sizes for the estimated parameters and posterior probability as calculated with the program Tracer v1.4[79] were higher than 1,000.

Nuclear introns primers, PCR and genotyping. Four nuclear introns, two found on the X-chromosome BGN[66,], PHKA2[67], the Y-Chromosome SMCY7[68] and one on chromosome 16 for *Homo sapiens* (ZP2, this study) were amplified and sequenced in at least 10 individuals per country (Thailand and Myanmar) (see Supplementary Table S8 for primers used). BGN, PHKA2 and SMCY7 were amplified using the same protocol as for mtDNA (see above), and ZP2 was amplified using the same protocol as NUMTs (see above). For these four introns, we identified five parsimoniously informative sites (one in each intron for BGN, SMCY7 and ZP2 and two in PHKA2). For the five SNPs, we designed custom TaqMan SNP assays (Applied Biosystems) and screened all the samples (n=659). Primer and probes sequences are reported on Supplementary Table S9. All reactions were conducted in 9 µl reaction volumes containing 1 µl of DNA , 4.5 µl 2X TaqMan® Genotyping Master Mix (Applied Biosystems), 0.225 µl 40X custom probe and 3.275 µl ddH2O. SNPs were genotyped on an ABI 7500 Fast Real-Time PCR System using a

pre-PCR read step of 60°C for 1 min, a denaturation step of 95°C for 10 min followed by 40 cycles (44

cycles for PHKA2-start) of 95°C for 15 s and 60°C for 1 min. A post-PCR read was done at 60°C for 1

min. SNPs were scored using the Allelic Discrimination Assay procedure in 7500 Software version 2.0.1

(Applied Biosystems).

Mismatch distributions analyses. The Thai populations' demographic history was examined using the

mismatch distribution of 462 mitochondrial DNA sequences[55]. Colony P (see Fig. 1) was not included in

the analysis because of the limited sampling (one individual). Episodes of population growth or decline

leave characteristic signatures in the distribution of nucleotide differences between pairs of individuals[80].

Mismatch distributions under demographic and spatial expansion scenarios differ from mismatch

distributions from stable populations, which typically present multimodal mismatch distributions[55,80].

Furthermore, contrary to stable populations, expanding populations present non-overlapping confidence

intervals of the mutation parameters θ_0 and θ_1, θ_0 and θ_1 being proportional to the effective population

size before and after expansion respectively. When the confidence interval of θ_1 does not overlap with the

confidence interval of θ_0, the population after expansion is significantly larger than the population before

expansion.

We first calculated, in Arlequin version 3.5[58], confidence intervals of the two mutation

parameters for simulated data sets of stable and expanding populations. These results supported an

expansion scenario for all populations which had non-overlapping 99% confidence intervals for θ_0 and θ_1.

Therefore, to identify whether the expansion was spatial or demographic, the observed mismatch

distribution was compared to mismatch distribution simulated under: a pure demographic expansion

model[55,80]; and a spatial expansion model[21,81]. A pure demographic expansion model assumes that a

stationary panmictic population has suddenly passed from a population size of N_0 to N_1. This scenario has

been shown to lead to star shaped gene genealogies[55], translating into an excess of rare mutations and into

unimodal mismatch distributions[80]. A model of spatial expansion assumes an initial panmictic population

with a limited distribution range. The population's distribution range increases over time and space,

leading to subdivided populations in the sense that individuals in geographic proximity are more likely to

mate with each other than with remote individuals[21]. The genetic signature under this scenario is similar

to a pure demographic expansion model when the number of migrants between the subdivided

populations is high. However, when the number of migrants is small, the mismatch distribution is

bimodal[21,81] and the first peak in the distribution is the result of comparison of identical sequences.

The raggedness index (Rag.) was used to estimate how well the model fitted the observed data.

Confidence intervals (CI) and P-values for model rejection were obtained by 10,000 parametric

bootstrapping. Calculations were carried out in Arlequin version 3.5[58] and graphs were generated with R

version 2.12.0[46].

Causal modelling. We used causal modelling on resemblance matrices[23,24] to investigate: (i) the

combination of variables driving genetic differentiation between colonies; and, (ii) the causal

relationships between these variables and genetic differentiation. To investigate the combination of

variables driving genetic differentiation, we identified *a priori* three variables that could influence (or be

influenced) by genetic distance: geographic distance, the presence of barriers and echolocation

difference. The diagnostic set of statistical tests of the seven possible organisational models including

these variables was then evaluated (see Supplementary Table S4). Only models where all of their

diagnostic set of statistics fit the observed data were considered as fully supported. All possible causal

relationships between variables present in the fully supported organisational model(s) were then

investigated (see Supplementary Table S5), except those where geographic distance would be a

dependant variable as geographic distance is not influenced by echolocation or genetic distance[23]. As

originally described[23,24], we used Mantel[82] and partial Mantel tests[83] to assess the support of the

organisational models. All tests (one-sided) were conducted with R version 2.12.0[84] using the package

'ecodist' version 1.1.4[59], and significance was assessed with 9999 permutations[85]. We used a sequential

Bonferroni technique to correct for multiple testing[86] considering an overall significance level of 0.05.

Each data set was first translated into a pairwise distance matrix that represented the difference between

each pair of colonies. Tests were carried out on 12 colonies for which we obtained echolocation (n=3958

calls in total) and genetic data (n=442 individuals).

A matrix of genetic distance between colonies was calculated using the F-statistics option in

Genepop version 4.0.6[53,87]. The locus CTC1 was excluded when calculating the genetic distance matrix as

it was shown to be under selection (see main text) and therefore, was not representative of neutral variation. Geographic Distance Matrix Generator version 1.2.2[88] was used to generate a geographic distance matrix between colonies from geographic coordinates. The absence of suitable habitat and roosts can represent a strong barrier to gene flow[89,90]. Therefore, the presence/absence of at least one barrier between two colonies was entered in a matrix as '1'= presence or '0'= absence. A barrier was set as present if the limestone formation was discontinuous between two colonies for more than 1 km. The 1 km threshold corresponds to the maximum distance from the cave *C. thonglongyai* has been shown to forage at[34]. Finally, we calculated a matrix of absolute difference of the mean echolocation call frequency between colonies as in Yoshino *et al.*[90]. As explained previously, to remove any potential bias caused by the non-independence between calls from the same recordings, we generated 10,000 data sets by randomly picking only one call per recording. We then calculated 10,000 matrices of absolute echolocation difference and performed the partial Mantel test for each new matrix. We therefore obtained 10,000 Mantel *r* and P-values from which we calculated the median value.

Estimates of dispersal distance. Given the very narrow and elongated distribution of *Craseonycteris* along the Kwae river valley[16], we considered the habitat as one-dimensional[91]. We then estimated the average axial dispersal distance (σ) using the following formula[91,92]: $b=1/(4D\sigma^2)$
where *b* is the slope of the regression of *Fst*/(1-*Fst*) *versus* distance and *D* the population density. We used two population density values here, the observed density ($258/7.6=33.95$ individuals/km^2, calculated after [16]) and an estimate of the effective density (taken as 1:10th of the observed density[93,94]). Although this estimate assumes a population in equilibrium, it was shown to be relatively robust to various scenarios of temporal and spatial fluctuations of demographic parameters[95,96]. The estimate was shown to be biased when demographic expansion occurred less than 100 generations ago[95], which is unlikely to be the case for *Craseonycteris* given estimates of time since expansion calculated based on mismatch distributions. Indeed, based on mismatch distributions dating ($\tau=2ut$)[80], a mutation rate of 2050% per million year (95% confidence interval: 1130-2680) would be required for the expansion to have occurred 100 generations ago. This value is well above the maximum values reported in any mammalian species

for any mtDNA fragment, including the D-loop alone[97,98], suggesting that the expansion occurred more than 100 generations ago.

Atmospheric attenuation calculations. The main parameters affecting atmospheric attenuation of sound in the air are relative humidity, temperature and frequency of the sound. To keep the detection distance of prey and objects constant, it is predicted that the peak frequency used in warmer places should be lower than that used in colder places; also, lower frequencies should be used in more humid places[31]. Relative humidity, temperature and frequency interact in a complex way to attenuate sound in the air[99] and therefore influence the maximum distance of prey detection[30]. We calculated sound attenuation for peak frequencies from each of the thirteen sites in Thailand according to formulas presented in Bazley[99]. Monthly averages of maximum values of relative humidity and temperature at 2 m were obtained from the EMP Climate database provided by the Center for Energy and Processes of Mines ParisTech/Armines (Resolution, 5 arcmin). The average peak frequency at each site was measured from recordings of free flying bats around caves' entrances (*cf.* "Echolocation Calls" paragraph in the Methods).

If the *Craseonycteris* change in frequency between the different colonies in Thailand was an adaptation to produce calls with similar attenuation throughout the range, we would expect to observe similar attenuation values for the different colonies. Also, these attenuation values should be less variable between sites than if all colonies had a similar peak frequency throughout Thailand. To check this, we calculated attenuation for the observed frequency at each site as well as for a theoretical constant frequency for all colonies (considered values of 74 kHz, 75.5 kHz and 77 kHz). Calculations were carried out in R version 2.12.0[46] and are presented in Figure 6.

Echolocation drift. We investigated the short-term and long-term population size of *C. thonglongyai* in Thailand with particular emphasis on the difference between the Northern/Central and Southern colonies. The actual colony size was estimated at 20 sites spread throughout the species range by counting individuals emerging from caves at dusk (see details in Puechmaille *et al.*[16]). The average colony size was then compared between the Northern/Central and Southern colonies using a one-sided Wilcoxon Mann-Whitney Rank Sum Test[100] implemented in the 'coin' package in R version 2.12.0[84]. We used the genetic

diversity found within colonies as a proxy for long-term effective population size. High genetic

diversities are associated with large effective population sizes and reduced drift whereas low genetic

diversities are associated with reduced population sizes and increased drift[25]. To test for differences in

genetic diversity between the Northern/Central and Southern colonies, we grouped colonies into two

groups corresponding to colonies situated north *versus* south of the echolocation break and compared

their genetic diversity at the nuclear and mitochondrial level in Fstat version 2.9.3.2[101] as described in the

"Genetic diversity indices**"** section in the Methods. For the two tests detailed above, the North-South

separation of colonies was identical to what is described in the legend of Figure 5. The same tests were

applied to compare genetic diversity between the Thai and Myanmar populations.

Selection tests. Individuals from different populations living in different environments often vary

genetically at key sites in the genome due to adaptation to different local conditions[26,102]. Low genetic

differentiation across populations may indicate balancing selection, whereas high genetic differentiation

suggests positive directional selection[103]. The 15 microsatellites from 462 individuals were tested for

selection based on the *Fst* outlier approach[27,104]. Colony P (see Fig. 1) was not included in the analysis

because of the limited sampling (one individual captured). This method evaluates the relationship

between *Fst* and H_E (expected heterozygosity) under the assumption of neutrality. We carried out the test

under an island model as implemented in LOSITAN[60] and under a hierarchical island model[105] as

implemented in Arlequin version 3.5. The first model has been tested under a wide variety of conditions,

including isolation by distance[27], and is robust to deviations from the theoretical model used[27,104,106],

while the second method has been specifically designed to model hierarchical population structure.

The tests for the island model were performed using the program LOSITAN under the infinite

alleles and stepwise mutation models. Significance was assessed by the proportion of simulated values

having higher (balancing selection) or lower (positive selection) *Fst* values than expected under neutrality

(α level set to 0.01). As recommended by Beaumont and Balding[106], the cut-off *P*-value for significance

was adjusted for the number of loci being tested using a sequential Bonferroni correction[86]. As both

models of mutation gave similar results, only results of the infinite allele model are presented. The

"Neutral mean *Fst*" option was used when running LOSITAN. This option allows the program to run

once to determine a first candidate subset of selected loci in order to remove them from the computation of the neutral Fst, which is generally a better approximation of the neutral Fst[27]. The option "Force mean Fst" was also used to get a simulated mean Fst close to the observed one found in the real dataset[60]. All other options or settings were left as default and 50,000 simulations were run.

For microsatellites under positive selection, we estimated the allelic and genotype frequency distribution across the different colonies and identified geographic areas (sites) where frequencies dramatically changed, potentially indicating areas subject to divergent selection. To confirm that the divergent selection was occurring within this geographic region we ran the selection tests for the colonies on either side of potential selection border.

The test for the hierarchical island model was run in Arlequin with 10 simulated groups, 100 demes per group and 10,000 simulations. Excoffier *et al*[105] showed that "*the splitting of existing groups is less detrimental than the lumping of those groups*" so we ran the test for different population structure with up to five groups. The population structures were defined according to the analysis of the first three axis of a PCA analysis on colonies (data not shown). The first structure involved two groups with all Southern *versus* Northern/Central colonies (as in Fig. 5b-c). The second structure was represented by three groups, keeping the Southern group from the previous structure and splitting the Northern group into two (A, B, D and E, F, G). The third structure was constituted of the two groups from the North (as detailed above), plus three groups resulting from the separation of the Southern colonies (1: K; 2: H, I, J and 3: L, M, N, O).

Supplementary References:

61 Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., & Saccone, C. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* **205**, 125-140 (1997).

62 Nikaido, M. *et al.* Maximum likelihood analysis of the complete mitochondrial genomes of Eutherians and a reevaluation of the phylogeny of bats and insectivores. *J. Mol. Evol.* **53**, 508-516 (2001).

63 Lin, Y.-H. & Penny, D. Implications for bat evolution from two new complete mitochondrial genomes. *Mol. Biol. Evol.* **18**, 684-688 (2001).

64 Lin, Y.-H. *et al.* Four new mitochondrial genomes and the increased stability of evolutionary trees of Mammals from improved taxon sampling. *Mol. Biol. Evol.* **19**, 2060-2070 (2002).

65 Pumo, D.E. *et al.* Complete mitochondrial genome of a Neotropical fruit bat, *Artibeus jamaicensis*, and a new hypothesis of the relationships of bats to other Eutherian Mammals. *J. Mol. Evol.* **47**, 709-717 (1998).

66 Lyons, L.A. *et al.* Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nat Genet* **15**, 47-56 (1997).

67 Murphy, W.J., Sun, S., Chen, Z.-Q., Pecon-Slattery, J., & O'Brien, S.J. Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid map. *Genome Res.* **9**, 1223-1230 (1999).

68 Hellborg, L. & Ellegren, H. Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. *Mol. Ecol.* **12**, 283-291 (2003).

69 Kumar, S., Tamura, K., & Nei, M. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **5**, 150-163 (2004).

70 Bensasson, D., Zhang, D.-X., Hartl, D.L., & Hewitt, G.M. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.* **16**, 314-321 (2001).

71 Triant, D.A. & deWoody, J.A. The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. *J. Mammal.* **88**, 908-920 (2007).

72 Lopez, J.V., Culver, M., Claiborne Stephens, J., Johnson, W.E., & O'Brien, S.J. Rates of nuclear and cytoplasmic mitochondrial DNA sequence divergence in Mammals. *Mol. Biol. Evol.* **14**, 277-286 (1997).

73 Zischler, H., Geisert, H., von Haesseler, A., & Pääbo, S. A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. *Nature* **378**, 489-492 (1995).

74 Dixon, C.J. A means of estimating the completeness of haplotype sampling using the Stirling probability distribution. *Mol. Ecol. Notes* **6**, 650-652 (2006).

75 Posada, D. & Crandall, K.A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818 (1998).

76 Drummond, A.J., Ho, S.Y.W., Rawlence, N., & Rambaut, A. A Rough Guide to BEAST 1.4, 2007.

77 Pesole, G., Gissi, C., De Chirico, A., & Saccone, C. Nucleotide substitution rate of Mammalian mitochondrial genomes. *J. Mol. Evol.* **48**, 427-434 (1999).

78 Laslett, D. & Canback, B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**, 172-175 (2008).

79 Rambaut, A. & Drummond, A. Tracer v1.4 MCMC Trace Analysis Package, Institute of Evolutionary Biology University of Edinburgh (Scotland) & Department of Computer Science University of Auckland (New Zealand) (2003-2007), http://beast.bio.ed.ac.uk/Tracer, 2007.

80 Rogers, A.R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552-569 (1992).

81 Excoffier, L. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol. Ecol.* **13**, 853-864 (2004).

82 Mantel, N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209-220 (1967).

83 Smouse, P.E., Long, J.C., & Sokal, R.R. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* **35** (1986).

84 Ihaka, R. & Gentleman, R. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299-314 (1996).

85 Jackson, D.A. & Somers, K.M. Are probability estimates from the permutation model of Mantel's test stable? *Can. J. Zoolog.* **67**, 766-769 (1989).

86 Rice, W.R. Analyzing tables of statistical tests. *Evolution* **43**, 223-225 (1989).

87 Raymond, M. & Rousset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248-249 (1995).

88 Ersts, P.J. Geographic Distance Matrix Generator, 2007.

89 Worthington Wilmer, J., Hall, L., Barratt, E., & Moritz, C. Genetic structure and male-mediated gene-flow in the ghost bat (*Macroderma gigas*). *Evolution* **53**, 1582-1591 (1999).

90 Yoshino, H., Armstrong, K.N., Izawa, M., Yokoyama, J., & Kawata, M. Genetic and acoustic population structuring in the Okinawa least horseshoe bat: are intercolony acoustic differences maintained by vertical maternal transmission? *Mol. Ecol.* **17**, 4978-4991 (2008).

91 Rousset, F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228 (1997).

92 Rousset, F. Genetic differentiation between individuals. *J. Evol. Biol.* **13**, 58-62 (2000).

93 Frankham, R. Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* **66**, 95-107 (1995).

94 Vucetich, J.A., Waite, T.A., & Nunney, L. Fluctuating population size and the ratio of effective to census population size. *Evolution* **51**, 2017-2021 (1997).

95   Leblois, R., Rousset, F., & Estoup, A. Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* **166**, 1081-1092 (2004).

96   Leblois, R., Estoup, A., & Rousset, F. Influence of mutational and sampling factors on the estimation of demographic parameters in a "continuous" population under isolation by distance. *Mol. Biol. Evol.* **20**, 491-502 (2003).

97   Petit, E., Excoffier, L., & Mayer, F. No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution* **53**, 1247-1258 (1999).

98   Santos, C. *et al.* Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores islands (Portugal). *Mol. Biol. Evol.* **22**, 1490-1505 (2005).

99   Bazley, E.N. National Physics Laboratory Acoustics Report 74, 1976.

100 Hollander, M. & Wolfe, D.A. *Nonparametric Statistical Methods*, 2nd Edition ed. (John Wiley & Sons, New York, 1999).

101 Goudet, J. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2), 2001.

102 Stinchcombe, J.R. & Hoekstra, H.E. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**, 158-170 (2007).

103 Wood, H.M., Grahame, J.W., Humphray, S., Rogers, J., & Butlin, R.K. Sequence differentiation in regions identified by a genome scan for local adaptation. *Mol. Ecol.* **17**, 3123-3135 (2008).

104 Beaumont, M.A. Adaptation and speciation: what can *Fst* tell us? *Trends Ecol. Evol.* **20**, 435-440 (2005).

105 Excoffier, L., Hofer, T., & Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285-298 (2009).

106 Beaumont, M.A. & Balding, D.J. Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* **13**, 969-980 (2004).

107 Ulanovsky, N. & Moss, C.F. What the bat's voice tells the bat's brain. *Proc. Natl Acad. Sci. USA* **105**, 8491-8498 (2008).

108 Neuweiler, G. Evolutionary aspects of bat echolocation. *J. Comp. Physiol. A* **189**, 245-256 (2003).

109 Kober, R. & Schnitzler, H.-U. Information in sonar echoes of fluttering insects available for echolocating bats. *J. Acoust. Soc. Am.* **87**, 882-896 (1990).

110 Roverud, R.C. A time window for distance information processing in the bats, *Noctilio albiventris* and *Rhinolophus rouxi* in *Animal sonar Processes and Performance*, edited by Paul E. Nachtigall & Patrick W. B. Moore (Plenum Press, New York, 1988), Vol. 156, pp. 513-517.

111 Bates, M.E., Stamper, S.A., & Simmons, J.A. Jamming avoidance response of big brow bats in target detection. *J. Exp. Biol.* **211**, 106-113 (2008).

112 Gillam, E.H., Ulanovsky, N., & McCracken, G.F. Rapid jamming avoidance in biosonar. *Proc. R. Soc. Lond. B* **274**, 651-660 (2007).

113 Ulanovsky, N., Fenton, M.B., Tsoar, A., & Korine, C. Dynamics of jamming avoidance in echolocating bats. *Proc. R. Soc. Lond. B* **271**, 1467-1475 (2004).

114 Habersetzer, J. Adaptive echolocation sounds in the bat *Rhinopoma hardwickei*. *J. Comp. Physiol. A* **144**, 559-566 (1981).

115 Fenton, M.B. *Communication in the Chiroptera*. (Indiana University Press, Bloomington, 1985).

116 Brigham, R.M., Cebek, J.E., & Hickley, M.B.C. Intraspecific variation in the echolocation calls of two species of insectivorous bats. *J. Mammal.* **70**, 426-428 (1989).

117 Obrist, M.K. Flexible bat echolocation: the influence of individual, habitat and conspecifics on sonar signal design. *Behav. Ecol. Sociobiol.* **36**, 207-219 (1995).

118 Fenton, M.B. & Ratcliffe, J.M. Eavesdropping on bats. *Nature* **429**, 612-613 (2004).

119 Russo, D. *et al.* Divergent echolocation call frequencies in insular rhinolophids (Chiroptera): a case of character displacement? *J. Biogeogr.* **34**, 2129-2138 (2007).

120 Jones, G. & Parijs, S.M.v. Bimodal echolocation in pipistrelle bats: are cryptic species present? *Proc. R. Soc. Lond. B* **251**, 119-125 (1993).

121 Podos, J. Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature* **409**, 185-188 (2001).

122 Voigt, C.C. *et al.* Songs, scents, and senses: sexual selection in the greater sac-winged bat, *Saccopteryx bilineata*. *J. Mammal.* **89**, 1401-1410 (2008).

123 Behr, O. *et al.* Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behav. Ecol.* **17**, 810-817 (2006).

124 Grillot, M.E. PhD dissertation, Auburn University, 2007.

125 Pfalzer, G. & Kusch, J. Structure and variability of bat social calls: implications for specificity and individual recognition. *J. Zool.* **261**, 21-33 (2003).